DBRE 輪読会 vol.02

2021/07/01

~サービスレベルマネジメント~

サービスレベルマネジメント

“ サービスがあるべき運用レベルを見定めること。
これがサービスの設計、ビルド、デプロイにあたって最初にすべきことです。 ”

この章で学ぶこと

- サービスの運用レベルについて、何をどのように定義するか
- 定義した運用レベルを満たしているか、どう測定・監視するか

目次

1. SLO とは
2. SLO の指標と定義
3. SLO に基づいた監視とレポート
4. まとめ

目次

1. **SLO とは**
2. SLO の指標と定義
3. SLO に基づいた監視とレポート
4. まとめ

1. SLO とは

- SLO (Service level Objective)
  - 設計及び運用に関して遵守すべき数値目標をまとめたもの
- SLA (Service Level Agreement)
  - SLA には SLO が特定のレベルを満たすことを約束する契約が規定されている

1. SLO とは

SLA の例)

月間稼働率	サービスクレジット率
99.0% 以上、99.99% 未満	10%
95.0% 以上、99.0% 未満	30%

参考: Amazon Compute サービスレベルアグリーメント

1. SLO とは

サービスレベルマネジメントは理解・設定が難しい

- 98%のユーザーは 99.99%の可用性、2%には 30%しか提供できない
- DB において、
  - 1 日分のデータが消失したが影響範囲が特定テーブルだけの場合
  - ユーザーがその損失に気づかない場合は?
- API サービスの SLO において、ユーザーの行う無効・不正なリクエストは、どう取り扱うべきか
- エラー発生率は全体の時間で平均して算出するのか、閾値を超えた数をカウントするのか

1. SLO とは

重要なのは、

その指標がユーザー体験、ひいてはユーザー満足度を反映

しているかどうか

目次

1. SLO とは
2. **SLO の指標と定義**
3. SLO に基づいた監視とレポート
4. まとめ

2. SLO の指標と定義

SLO はサービスが持つ特性によって変わる。
ただし、SLO の中心として考えるべきはあくまでもユーザー

- 代表的な 3 つの指標
  - レイテンシ
    - 1 リクエストに対するレスポンスにかかる時間
  - 可用性
    - システムが利用可能である状態を、%で示したもの
  - スループット
    - 一定時間において正常に処理されたリクエストの数

2.1. レイテンシ

別名:レスポンスタイム、RTT(ラウンドトリップタイム)

- 1 リクエストに対するレスポンスにかかる時間を表す
- エンドツーエンドで測定するのがもっとも望ましい

レイテンシが 100 ミリ秒を超えると、ビジネスに多くの影響を及ぼす

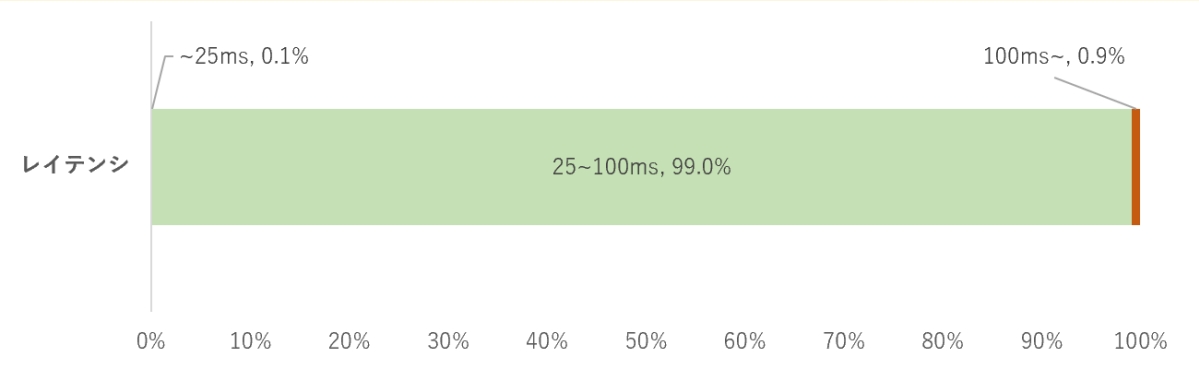
- Amazon: サイト表示が 0.1 秒遅くなると、売り上げが 1%減少し、1 秒高速化すると 10%の売上が向上する
- Google: サイト表示が 0.5 秒遅くなると、検索数が 20%減少する

2.1. レイテンシ

- ネットワークの通信速度は 0 にはならない
- レイテンシの実測値のうち、極端に高いものは取り除く

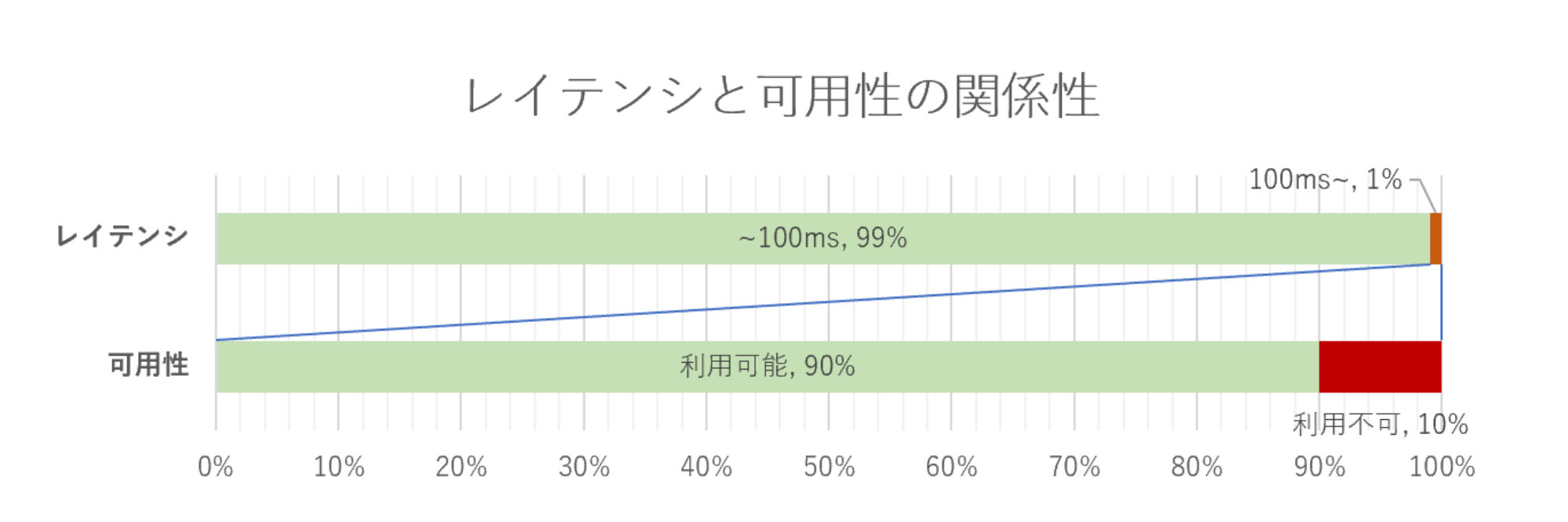
これにより SLO は、以下のようになる

リクエストの 99%において、レイテンシは 25 ミリ秒から 100 ミリ秒の間でなければならない



2.2. 可用性

- システムが利用可能である状態を、時間軸に対して%で示したもの
- レイテンシと合わせて評価されるべき指標



2.2. 可用性

SLO の具体的な表記は以下のようになる

- 数週間の時間軸で捉えた場合、それぞれの 1 週間あたりの可用性を計算して、その平均が 99.9%であること
- 障害が発生した場合、10.08 分以内に復旧すること
- ダウンタイムとは全体のユーザー数の 5%以上に影響を与える事象
- 1 年に 1 回は 4 時間のダウンタイムを許容。ただし、以下を満たす
  - 2 週間前にユーザーへ通知
  - 1 度にユーザー数の 10%以上に影響を与えない

2.3. スループット

一定時間において正常に処理されたリクエストの割合のこと
サービスが対応可能な最大値を設定する

例) システムが 1 秒間に 50 件のレスポンスを返せる場合は

50 件/秒

と表記

2.4. 耐久性

ストレージに対して一定の成功率で書き込みができるかどうか

例: Amazon S3 の耐久性

2.5. 費用対効果

費用は、1 ページビューや 1 購読数、購入 1 件といった行動に対して効果として測定される

何のために費用をかけたのか、認識することが重要

- ex.
  - EC サイト: トランザクション数
  - コンテンツプロバイダ: ページビュー
  - オンラインサービス: ユーザー数

2.6. サービスを運用するために求められること

- 新しいサービスの場合
  - 運用の指針となる SLO を定義
  - SLO として参照されるメトリクスに対して、
    目標及び実測値を適切に評価できる監視システムを設定

2.6. サービスを運用するために求められること

- 既存サービスの場合
  - サービスの過去と現在の状態を踏まえ、今までに達成した SLO と違反した SLO に対して、定期的なレビューを行う
- サービスに付随する問題
  - サービスレベルに影響をあたえうる不安要素を整理し、特定の不具合に対する回避策や修正度合いを検討する

2.7. SLO を定義する際に注意すること

- **ユーザー中心**
  - ユーザーにとって最も失われてはならないものから、SLO を組み立てる
- あれもこれもと欲張らない
  - 注目すべき指標のリストは、ダッシュボード 1 ページに収まる簡潔なものとする
    - たくさんあると、大切な指標を見逃してしまう
- SLO は定期的に見直す
  - ビジネスの段階によって、SLO に求められる内容が変わる

目次

1. SLO とは
2. SLO の指標と定義
3. **SLO に基づいた監視とレポート**
4. まとめ

3. SLO に基づいた監視とレポート

重要なのは
達成を妨げる潜在的なリスクを洗い出しその対策をすること

SLO 遵守のために、アラートを受ける前から不測の事態に対応する

そのために、以下の 3 つを考える

- 収集と分析の自動化
- 問題発生時のアラートの対応とその後のレポート
- 分析結果の視覚化

3.1. 可用性の監視

- RUM(Real User Monitoring)
  - ユーザーからのリクエストに対するエラー発生率
  - 累積したデータから近い将来のエラー発生率を予測する
  - 上記から週あたりのダウンタイムを超過しないか判断する
  - 潜在的な問題に対し、データドリブンな観点から目を光らせ、深刻な障害を未然に防ぐ

3.1. 可用性の監視

- 定点モニタリング (Synthetic Monitoring)
  - 故意に作成したデータセットを用いたテストを走らせる
  - カバレッジ計測にその効果を発揮する
    - チューニングされたクエリとカバレッジによって、異なる時間、地域の測定を行う

3.2. レイテンシの監視

> リクエストの 99%において、レイテンシは 25 ミリ秒から 100 ミリ秒の間でなければならない

レイテンシの SLO が上記の場合、

- HTTP のリクエストログを、時系列のデータとして保存する
- 上位 1%のデータを分析から除外する
- 1 秒毎におけるデータの 99%が 100 ミリ秒を超えていた場合、SLO 違反としてダウンタイムを計上する

3.3. スループットの監視

スループットの監視には以下が必要。

- 測定
- 収集
- 可用性とレイテンシの SLO と絡めたレビュー

秒単位でトランザクション数を記録しておく。

3.4. 費用対効果の監視

- ストレージ、CPU、メモリ等々で
  どれだけコストが発生しているのかを読み解く
- サービスを運営する人件費も考慮し、定期的に見直す

3.4. 費用対効果の監視

> サービスが生み出した価値と、それにかかる費用をエンジニアが理解することで、技術的な観点から無駄を省き、燃費のよいアーキテクチャを設計する動機が生まれ、結果的に効率のよいコスト削減ができるようになるでしょう。

4. まとめ

> SLO を定義管理することは、インフラストラクチャを設計し運用する上での要石です。
> サービス対するすべての施策は、SLO を遵守するための手段に過ぎません。
> SLO が日々の活動全ての基礎なのです。

参考

- SRE の基本（2021 年版）: SLI、SLA、SLO の比較 | Google Cloud Blog
- Maintain SLO ～俺たちの SLO はこれからだ!～ | メルカリエンジニアリング
- レイテンシとは：定義から計測サイトまで用語にまつわるトピックを解説
- Web の表示が〇秒遅くなると ×× まとめ。Web パフォーマンスの重要性を示すフレーズ集 - アイデアマンズブログ
- 担当マイクロサービスの SLI/SLO を見直そうと思ったんだ - エムスリーテックブログ